

Q65793

DIALOG(R)File 351:Derwent WPI  
(c) 2001 Derwent Info Ltd. All rts. reserv.

012011374 \*\*Image available\*\*

WPI Acc No: 1998-428284/199836

XRPX Acc No: N98-334264

**Distributed data storage system for intranet information system - has distributed storage server driven units and assignment manager distributing over total virtual storage.**

Patent Assignee: BULL SA (SELA )

Inventor: PEPING J

Number of Countries: 026 Number of Patents: 005

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week
WO 9833297	A2	19980730	WO 98FR110	A	19980122	199836 B
FR 2759517	A1	19980814	FR 97757	A	19970124	199838
EP 866410	A1	19980923	EP 98400133	A	19980122	199842
JP 11508078	W	19990713	JP 98531670	A	19980122	199938
			WO 98FR110	A	19980122	
US 6185655	B1	20010206	WO 98FR110	A	19980122	200109
			US 98142859	A	19980916	

Priority Applications (No Type Date): FR 97757 A 19970124

Cited Patents: No-SR.Pub

Patent Details:

Patent No	Kind	Lan Pg	Main IPC	Filing Notes
-----------	------	--------	----------	--------------

WO 9833297	A2 F	26	H04L-012/00	
------------	------	----	-------------	--

Designated States (National): JP US

FR 2759517	A1		H04L-012/28	
------------	----	--	-------------	--

EP 866410	A1 F		G06F-017/30	
-----------	------	--	-------------	--

Designated States (Regional): AL AT BE CH DE DK ES FI FR GB GR IE IT LI

LT LU LV MC MK NL PT RO SE SI

JP 11508078	W	32	G06F-015/16	Based on patent WO 9833297
-------------	---	----	-------------	----------------------------

US 6185655	B1		G06F-012/00	Based on patent WO 9833297
------------	----	--	-------------	----------------------------

Abstract (Basic): WO 9833297 A

The electronic data processing system has several mechanisms for storage of distributed data, and at least one storage server. There is a distribution mechanism for each server, with a virtual memory space, which is at least equal to the cumulative capacity of all the distributed data storage areas.

Preferably the distribution mechanism consists of a series of first data processing sections with recording programmes, which manage each part of the data storage mechanisms. Preferably there are second data processing mechanisms, associated with a centralised administration section, and associated with a descriptive database of a configuration arrangement for the data storage mechanism.

ADVANTAGE - Information system is adaptable to future technologies without needing extensive modification.

Dwg. 4/8

Title Terms: DISTRIBUTE; DATA; STORAGE; SYSTEM; INFORMATION; SYSTEM;  
DISTRIBUTE; STORAGE; SERVE; DRIVE; UNIT; ASSIGN; MANAGE; DISTRIBUTE;  
TOTAL; VIRTUAL; STORAGE

Derwent Class: T01; W01

International Patent Class (Main): G06F-012/00; G06F-015/16; G06F-017/30;

**THIS PAGE BLANK (USPTO)**

H04L-012/00; H04L-012/28  
International Patent Class (Additional): G06F-012/00  
File Segment: EPI  
Manual Codes (EPI/S-X): T01-H07C5E; T01-M02A1C; W01-A06B7; W01-A06E2A

**THIS PAGE BLANK (USPTO)**

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表平11-508078

(43) 公表日 平成11年(1999) 7月13日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 15/16

識別記号

4 3 0

3 7 0

F I

G 0 6 F 15/16

4 3 0 B

3 7 0 N

審査請求 有 予備審査請求 未請求(全 32 頁)

(21) 出願番号 特願平10-531670  
 (86) (22) 出願日 平成10年(1998) 1月22日  
 (85) 翻訳文提出日 平成10年(1998) 9月22日  
 (86) 国際出願番号 P C T / F R 9 8 / 0 0 1 1 0  
 (87) 国際公開番号 W O 9 8 / 3 3 2 9 7  
 (87) 国際公開日 平成10年(1998) 7月30日/  
 (31) 優先権主張番号 9 7 / 0 0 7 5 7  
 (32) 優先日 1997年1月24日  
 (33) 優先権主張国 フランス (F R)  
 (81) 指定国 J P, U S

(71) 出願人 ブル・エス・アー  
 フランス国、エフー78430・ループシエンヌ、ルート・ドウ・ベルサイユ、68  
 (72) 発明者 ジャック、ベパン  
 フランス国、95150・タベルニー、リュ・ビクトール・バツシュ、72  
 (74) 代理人 弁理士 川口 義雄 (外2名)

(54) 【発明の名称】 分散型データ記憶装置を有する情報処理システム

(57) 【要約】

本発明による情報処理システム (1) においては、データの各記憶資源 ( $D_1$  ないし  $D_3$ 、 $F D_6$ 、 $T L_4$ 、 $S T O$ ) が、分散形管理装置 ( $D S M_1$  ないし  $D S M_6$ ) の制御下にある、サーバ ( $S_1$  ないし  $S_3$ ) に接続され、あるいは接続されない。この分散形管理装置は、集中形管理装置 ( $N S M$ ) から、外部記憶資源の全部または一部及びローカル記憶装置を有する仮想メモリ空間をダイナミックに割当ててくるコマンドプログラムを受取る。情報処理システムに適用可能である。

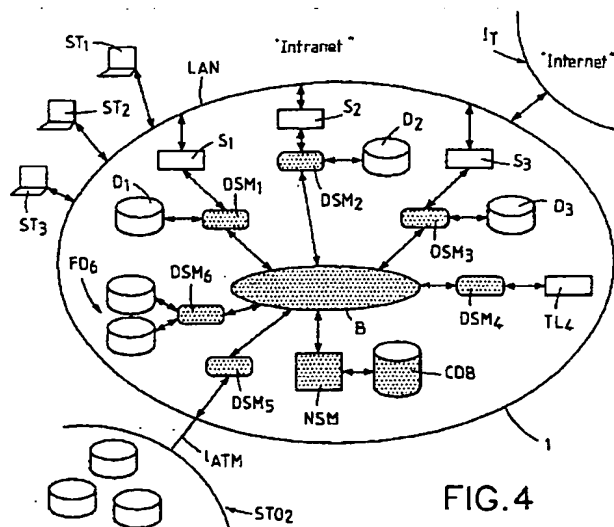


FIG. 4

## 【特許請求の範囲】

1. 複数の分散形データ記憶手段 ( $D_1-D_3$ 、 $FD_6$ 、 $TL_4$ 、 $STO_e$ ) 及びデータを記憶するための少なくとも一つのサーバ ( $S_1-S_3$ ) を備えた情報処理システム (1) であって、前記分散形データの前記記憶手段 ( $D_1-D_3$ 、 $FD_6$ 、 $TL_4$ 、 $STO_e$ ) の全部または一部の累積容量に少なくとも等しい範囲を有する仮想メモリ空間 ( $D'_{XYZ}$ ) を、前記データ記憶サーバ ( $S_1-S_3$ ) の各々に割当てて手段を備えていることを特徴とする情報処理システム。

2. 前記割当て手段が、分散形管理手段と呼ばれる内部記憶プログラム型の一連の最初のデータ処理手段によって構成され、各記憶手段は、前記データ記憶手段 ( $D_1-D_3$ 、 $FD_6$ 、 $TL_4$ 、 $STO_e$ ) の一部を管理することを特徴とする請求項 1 に記載のシステム。

3. 更に、少なくとも前記データ記憶手段 ( $D_1-D_3$ 、 $FD_6$ 、 $TL_4$ 、 $STO_e$ ) の構成を記述するデータのデータベース (CDB) に接続する集中形管理手段と呼ばれる第二のデータ処理手段 (NSM) を備え、該第二のデータ処理手段は、前

記分散形管理手段 ( $DSM_1-DSM_6$ ) が記憶サーバ ( $S_1-S_3$ ) に接続された時に、前記ダウンロードされたプログラムの制御の下に、前記仮想メモリ空間 ( $D'_{XYZ}$ ) を該記憶サーバ ( $S_1-S_3$ ) に割当てるように、前記データから前記プログラムを作成し、前記分散形管理手段 ( $DSM_1-DSM_6$ ) にダウンロードする手段を備えていることを特徴とする請求項 1 または 2 に記載のシステム。

4. 前記データ記憶手段 ( $D_1-D_3$ 、 $FD_6$ 、 $TL_4$ 、 $STO_e$ ) の少なくとも一部は、メモリ空間が、前記データ記憶サーバ ( $S_1-S_3$ ) によって送信される書込み-読取りタイプの入/出力命令を有するプロトコルによってアドレス指定可能であるようなローカルデータ記憶手段を形成するために、前記分散形管理手段 ( $DSM_1-DSM_6$ ) のいずれか一つを介して、データの記憶サーバ ( $S_1-S_3$ ) に接続されており、更に、前記ローカルデータ記憶手段が、前記仮想メモリ空間 ( $D'_{XYZ}$ ) の第一の区分を形成し、前記データ記憶サーバにおける外部記憶手段が、この仮想メモリ空間 ( $D'_{XYZ}$ ) の少なくとも第二の区分を形成す

ることを特徴とする請求項1から3のいずれか一項に記載のシステム。

5. 多数の入／出力ポートを備えており、書込み－読取りタイプの入／出力命令を含む多数のプロトコルの伝送を許可するタイプ的高速バス（B）を有し、前記分散形管理手段（ $DSM_1$ － $DSM_6$ ）及び前記集中形管理手段（NSM）が、バス（B）によって互いに接続され、前記データ記憶サーバ（ $S_1$ － $S_3$ ）によって送信された前記入／出力命令が、前記データ記憶サーバ（ $S_1$ － $S_3$ ）に割り当てられた前記仮想メモリ空間（ $D'_{XYZ}$ ）の集合を直接的にアドレス指定可能にするように、前記命令が、これらのどちらの区分に関するかに応じて、ダウンロードされた前記プログラムのコマンドの下に、仮想メモリ空間（ $D'_{XYZ}$ ）の前記第一区分を形成するローカル記憶手段に向かって、あるいは、バス（B）を介して、仮想メモリ空間（ $D'_{XYZ}$ ）の追加区分の一つを形成する外部記憶手段に接続された分散形管理手段（ $DSM_1$ － $DSM_6$ ）に向かって、前記分散形管理手段（ $DSM_1$ － $DSM_6$ ）によって経路付けされることを特徴とする請求項1から4のいずれか一項に記載のシステム。

6. 前記バス（B）を構成する伝送媒体が、モノモード光ファイバであり、前記ポートが、他のすべての管理手段に接続さ

れた管理手段の各々を起点とする直接的な伝送通信路をつくりだすことができるように、多数の交換器を備えていることを特徴とする請求項5に記載のシステム。

7. 前記記憶手段が、磁気ディスク装置（ $D_1$ － $D_3$ 、 $FD_6$ ）、磁気テープまたは磁気カートリッジ装置（ $TL_4$ ）、あるいは光ディスクを備えていることを特徴とする請求項1から6のいずれか一項に記載のシステム。

8. 前記データ記憶手段の少なくとも一つ（ $STO_e$ ）が、遠隔サイトに配置され、このサイトと前記情報処理システム（1）との間の伝送が、非同期モードにおいて高速回線（ $I_{ATM}$ ）によって行われ、更に前記回線（ $I_{ATM}$ ）が、前記分散形管理手段（ $DSM_5$ ）の一つに接続され、これらの手段自体が前記バス（B）に接続されていることを特徴とする請求項1から7のいずれか一項に記載のシ

ステム。

9. 前記集中形管理手段 (NSM) によって作成されるプログラムが、前記分散形管理手段 (DSM<sub>1</sub>—DSM<sub>6</sub>) に選択的にダウンロードされ、前記バス (B) によって伝送されるように構成され、該プログラムが、ダウンロードの際に自動実行できるようにする所定の言語で書込まれ、更に、前記集中形管

理手段 (NSM) が、時間とともに変化するパラメータと前記情報処理システム (1) によって行われる処理に応じて、前記ダウンロードをダイナミックに管理すること特徴とする請求項 1 から 8 のいずれか一項に記載のシステム。

10. 前記ダウンロードされたプログラムが、各記憶サーバ (S<sub>1</sub>—S<sub>3</sub>) に割当てられた仮想メモリ空間 (D' <sub>XYZ</sub>) の管理プログラムを含むことを特徴とする請求項 9 に記載のシステム。

11. 前記データ記憶手段 (D<sub>1</sub>—D<sub>3</sub>、FD<sub>6</sub>、TL<sub>4</sub>、STO<sub>e</sub>) が異種性であり、さまざまなモード及び／又はプロトコルに従って作動し、前記ダウンロードされたプログラムは、各記憶サーバ (S<sub>1</sub>—S<sub>3</sub>) が、固有のプロトコルを利用しながら割当てられた前記仮想メモリ空間 (D' <sub>XYZ</sub>) にアクセスするように、モード及び／又はプロトコルの翻訳プログラムを含むことを特徴とする請求項 9 に記載のシステム。

12. 前記ダウンロードされたプログラムが、前記分散形データ記憶手段 (D<sub>1</sub>—D<sub>3</sub>、FD<sub>6</sub>、TL<sub>4</sub>、STO<sub>e</sub>) の中で、所定の階層に従った、データの自動保存プログラムを含むことを特徴とする請求項 9 に記載のシステム。

13. 前記ダウンロードされたプログラムが、所定の冗長構成に従った、データの記憶によって該データをバックアップするプログラムを含むことを特徴とする請求項 9 に記載のシステム。

14. 前記データ記憶手段の少なくとも一つ (STO<sub>e</sub>) が、遠隔サイトに配置され、該サイトと前記情報処理システムとの間の伝送が、非同期モードにおいて、高速回線 (I<sub>ATM</sub>) によって行われ、前記ダウンロードされたプログラムが、データの伝送プロトコルの翻訳プログラムを含み、これらのプログラムが、非

同期モードにおいて、前記バス（B）と前記回線（I<sub>ATM</sub>）との間のインターフェースを形成する分散形管理手段（DSM<sub>5</sub>）にダウンロードされることを特徴とする請求項9に記載のシステム。

15. 前記分散形管理手段（DSM<sub>1</sub>－DSM<sub>6</sub>）の少なくとも一部が、データ記憶サーバ（S<sub>1</sub>－S<sub>3</sub>）または、論理装置を形成するように該分散形管理手段が接続される前記分散形データ記憶手段（D<sub>1</sub>－D<sub>3</sub>、FD<sub>6</sub>、TL<sub>4</sub>、STO<sub>e</sub>）の一部をなすことを特徴とする請求項1から13のいずれか一項に記載のシステム。

16. 追加データ処理装置（ST<sub>1</sub>－ST<sub>3</sub>）と、ローカルネットワーク（LAN）と、外部データ伝送ネットワーク（I<sub>T</sub>）への前記情報処理システム（1）の接続手段とを備え、前記追加装置（ST<sub>1</sub>－ST<sub>3</sub>）と前記接続手段と前記データ記憶サーバ（S<sub>1</sub>－S<sub>3</sub>）とが、前記ローカルネットワーク（LAN）に接続されることを特徴とする請求項1から15のいずれか一項に記載のシステム。

17. 前記バス（B）と前記ローカルネットワーク（LAN）が、単一のデータ伝送ネットワーク（B）に組み込まれ、前記追加装置（ST<sub>1</sub>－ST<sub>3</sub>）が、前記バス（B）に接続され、前記バス（B）を介して通信プロトコルに従って互いに通信することを特徴とする請求項16に記載のシステム。

## 【発明の詳細な説明】

### 分散形データ記憶装置を有する情報処理システム

本発明は、分散形データ記憶装置を有する情報処理システムに関する。

本発明は、ネットワークアーキテクチャ型データ処理情報処理システムに適用され、特に、企業または組織に普及している“INTRANET”と呼ばれるタイプのネットワークアーキテクチャに適用される。

一般に、企業または会社の強さのキーファクタの一つは、保持する情報に直接的に依存する。この“情報”という用語は、最も広範な意味を含むものでなければならない。従って、会社内部の情報（製品の販売価格、製造品のレンジ等々）または会社の外部からの情報（技術や広告についての競争相手に関するさまざまなデータ等々）が問題となる。

従来の技術においては、これらの情報の処理は、必然的に、高性能且つ複雑になりつつある情報処理システムに依存している。ハードウェア、特に大容量記憶装置（ハードディスク、磁気テープまたはカートリッジ装置、光ディスク）の価格の急落

によって、ローカルあるいは遠隔のサイトにおいて、次第に多くのデータを記憶することが可能になってきている。

また、データ処理システムは、次第にネットワークに統合されつつある。それらネットワークの中では、特に、家庭用のシンプルなマイクロコンピュータを含む世界中に散在している数百万台のコンピュータ間の対話を可能にする“インターネット”ネットワークに注目しなければならない。

同じように、企業は、それぞれに固有の一つまたは複数のサイトのさまざまな情報処理資源を互いにつなぐ“イントラネット”と呼ばれる特有のローカルネットワークを利用している。

このことから、企業にとっては、増大する入力情報の流れを掌握し、特にそれらの情報を、できるだけ効果的にまたできるだけ低コストでアクセス可能にするとともに、情報の移動や管理を行える“場所”に記憶させることが緊急に必要となっている。

それらのデータはまた、最も広範な言葉の意味において、保護されなければならない。一般に、使用可能性、無傷性、秘密性の略語“D. I. C.”で知られるコンセプトが利用されている。

実際に、システムの故障に対してであれ、悪意のある行為に対してであれ、データの無傷性を確保することが必要である。また、秘密にすべきデータもあるので、少なくとも、許可されたユーザにだけ限定されたアクセスに対する“予約”手段を備えることが必要である。更に、データはできるだけ使用可能でなければならない。これは、特に、ハードウェアの故障やソフトウェアのエラーに対処するために、記憶装置の中に保護手段またはある程度の冗長度をもたせることを意味する。

更に、記憶システムの選択がいったん行われると、そのシステムの永続性が保証されることも必要である。特に、システムの大幅な手直しを必要とせずに、将来のテクノロジーについて考慮することができなければならない。

本発明の目的は、上述の必要性を満たすことができるシステムを提供することである。

そのために、本発明は、情報処理システムの中に分配された記憶資源の一般化された仮想アドレス指定を用いて、データの分散形記憶を行なうことを含む。

従って、本発明は、複数の分散形データ記憶手段と、これらのデータを格納するための少なくとも一つのサーバを備える情

報処理システムであって、分散形データ記憶手段の全部または一部の累積容量に少なくとも等しい容量を有する仮想メモリ空間を、データ記憶サーバの各々に割当てて手段を備えていることを特徴とする情報処理システムを対象とする。

本発明の好ましい変形実施形態によれば、システムは、主に、集中形管理手段と、情報処理システムの少なくとも一つの記憶資源に結び付けられた分散形管理手段と、これらの分散形管理手段を互いに接続し且つ集中形管理手段に接続する高速通信バスとを備えている。

添付の図面を参照して、以下に詳細に説明することによって、本発明の他の特

性及び利点がより明らかになるであろう。

－第1図は、従来の技術によるネットワークタイプのデータ処理アーキテクチャの一例を概略的に示す図である。

－第2図は、このようなネットワークの二つのサーバ間のデータ交換を示すタイムチャートである。

－第3a図及び第3b図は、本発明によって動作する基本的アーキテクチャを概略的に示す図である。

－第4図は、本発明によるネットワークタイプの情報処理システム全体のアーキテクチャの実施例を概略的に示す図である。

－第5図は、第4図によるシステムの中で使用されるバスの形態を概略的に示す図である。

－第6図は、第4図の情報処理システムの特定データを処理する二つの装置間のデータ交換を示す図である。

－第7図は、第6図の二つの装置間のデータ交換を示すタイムチャートである。

。

－第8図は、第4図のアーキテクチャによるシステムの変形例を示す図である。

。

第1図は、従来の技術によるネットワークタイプの情報処理システムのアーキテクチャの一例を示している。図面を簡略化するために、ここには“LAN”（“Local Area Network”）タイプのローカルネットワークに接続された二つの装置しか示されていない。このネットワークは、バス（例えば“ETHERNET－イーサネット”方式）またはリングタイプ（例えば“TOKEN RING－トークンリング”方式）といったあらゆるタイプまたはあらゆる構成のものとすることができる。第1図においては、それぞれ $S_A$ 及び $S_B$ の二つのサーバが対象になると仮定される。サーバは $S_A$ は、ワークステーションまたはシンプルなマイクロコンピュータに置き

換えることもできる。これらのサーバ $S_A$ 及び $S_B$ は、バスの性質によって異なる

従来のインターフェース  $I_A$  及び  $I_B$  を介して、LANタイプのパスに物理的に接続される。これらインターフェースは、特に、バッファメモリ手段と、特に、宛先のアドレスを認識することができる符号化及び複合手段を備えている。各サーバは、入／出力チャンネル、それぞれ  $I/O_A$  または  $I/O_B$  を用いて対応するサーバに接続されたハードディスク  $D_A$  または  $D_B$  によって表わされている大容量メモリ手段に結び付けられると仮定される。以下の説明においては、このネットワークを略語 LAN で示す。

二種類の主要なデータ交換方法が存在する。

ローカルでは、読取りまたは書込みにおいて、その固有ディスク  $D_A$  のメモリアドレスにアクセスするデータ処理装置、例えばサーバ  $S_A$  は、ディスクの読取り／書込みヘッドに直接作用する基本命令を使用する。これらの基本命令は、“ $I/O$  チャンネル” という名称で知られる“入／出力”タイプである。また、使用されるプロトコルは“メモリアドレス指定”タイプのものである。アクセスはほとんど瞬間的なものである。アクセスは、ディスク装置の特性（平均アクセス速度、スループット）

及び使用されるモードによって左右される。使用されるモードは例えば、従来タイプ及び“Small Computer System Interface”の略称“SCSI”という名称の下に（ANSI や ISO と同じように）標準化されたタイプである。第1図には、このデータ交換モードによる、ディスク  $D_A$  からのファイル  $F_A$  の読取りが示されている。このファイル  $F_A$  に関連するデータは、割付表の中に記録された一定のアドレスで、ディスク  $D_A$  のトラック及びセクタにおいて直接読取られる。

反対に、データを読取る場合には、例えば、外部記憶装置に記憶されたファイル  $F_B$ 、この場合にはサーバ  $S_B$  に結びついたディスク  $D_B$  を読取る時には、LAN ネットワークと、サーバ  $S_A$  及び  $S_B$  と、インターフェース  $I_A$  及び  $I_B$  を経由する必要がある。サーバ  $S_A$  は、もはやハードディスク  $D_B$  のヘッドを直接制御することができない。データの交換は、メッセージを用い、使用されるローカルネットワークのタイプに固有の特定プロトコルを利用するデータパケットによる通信

モードに従って行われる。

第 2 図のタイムチャートは、LAN ネットワークによるこの

ようなデータ交換を概略的に示している。基本的なデータパケットの交換に必要な全時間は、少なくとも時間間隔  $T$  に等しくなる。この間隔  $T$  は、基本的な時間間隔  $T_1$  ないし  $T_5$  に分割される。任意の瞬間  $t = 0$  と  $t = t_1$  の間の時間間隔  $T_1$  は、サーバ  $S_B$  における処理時間に対応する。瞬間  $t = t_1$  と  $t = t_2$  の間の時間間隔  $T_2$  は、使用される特定の通信プロトコルと、ネットワークを支配する瞬間的伝送条件（ロード、接続された装置数、等々）に起因する時間のロスに対応する。瞬間  $t = t_2$  から  $t = t_3$  の間の時間間隔  $T_3$  は、伝送時間を表わす。この時間は、主に、接続された二つのステーション間の距離（この例においてはサーバ  $S_A$  及び  $S_B$  の隔たり）、回線のスループット、更にそれらより度合いは小さいが、回線の物理的性質（信号の伝播速度）によって異なる。更に、瞬間  $t = t_3$  と  $t = t_4$  の間の時間間隔  $T_4$  は、通信プロトコルの（到着への）貢献度を表わしている。更に、瞬間  $t = t_4$  と  $t = t_5$  の間の時間間隔  $T_5$  は、到着時の処理時間を表わしている。

この交換モードが、先行のモードよりはるかに遅いことは容易に理解できる。有効なデータ（例えば、ファイル  $F_B$  に関連するデータ）の他に、追加データ、特にプロトコルに固有のデ

ータ及び（オリジン及びターゲット）アドレスデータが伝送されなければならない。エラーを処理するためには、冗長データ（パリティ、エラーの検出及び／又は補正コード）を伝送することが必要となる。更に、場合によって、バスやリンクでの考えられる競合や衝突を考慮する必要がある。これらの問題を解消するために、“Carrier Sense Multiple Access / Collision Detection” または“搬送波探知多重アクセス / 衝突検出”の“CSMA / CD”のような特別な方法が提案されている。これらの方法は、規格 IEEE 802.3 の勧告の対象となっている。これらの現象はまた、伝送の平均時間を増やす。

本発明の主要な特徴によれば、情報処理システムのさまざまな記憶資源における情報データの分散形記憶が利用される。そのために、各サーバには、その固有のディスク（設けられている場合）及び外部ディスク及び／又は他のデータ記憶資源を含む大容量の仮想メモリ空間が割当てられる。従って、サーバは、上述した“I/Oチャンネル”タイプの命令によって、付与された仮想メモリ空間全体を直接的にアドレス指定することができる。換言すれば、サーバは、仮想読取り／書込みヘッドを制御

する。

第3 a 図及び第3 b 図は、本発明による基本的アーキテクチャを概略的に表わしている。

第3 a 図では、一方ではLANローカルネットワークに接続することかできる（第1 図）サーバ $S_X$ が、先の例と同じように、データのローカル記憶装置、例えば $D_X$ ディスクに結び付けられる。本発明の重要な特徴によれば、サーバは、本発明によるシステムの分散形管理手段を構成する物理的または論理的装置 $DSM_X$ を通してディスクと通信する。

第一の場合、装置 $DSM_X$ が物理的タイプの時には、入／出力チャンネル $I/O_X$ によってディスク $D_X$ に、また以下に詳述する特有バスBのシステムによって、外部ディスク、例えばディスク $D_Y$ 及び $D_Z$ に物理的に接続される記録プログラム型データの処理回路を有する。このバスBは、上記のディスク $D_Y$ 及び $D_Z$ にも同じように“I/Oチャンネル”タイプの命令を伝送することができる。ここでは“インテリジェントエージェント”と呼ばれるタイプの物理装置 $DSM_X$ の中に記録された特定のプログラムが、この特定のジョブを可能にする。限定的でない例として、装置 $DSM_X$ は、“UNIX”（登録商標）オ

ペレーションシステムの下で作動する中間ワークステーションによって構成することができる。

第二の場合、装置 $DSM_X$ が論理タイプの時には、上記のプログラムは、サーバ $S_X$ のランダムアクセスメモリ（RAM）の中に直接記録される。さまざまな

ディスク $D_X$ ないし $D_Z$ に物理的に接続されるのはこのサーバ $S_X$ であり、サーバは、固有の役割と分散形管理手段( $DSM_X$ )の役割を果たす。従って、管理手段はワークステーション $S_X$ に組込まれた純粹に論理的なユニットとなる。

池の変形実施形態においては、論理ユニットは、ディスクを使用する場合には、例えばコントローラの中で記憶資源に結びつけられた電子回路の中に直接組込むことができる。

第3b図によってより詳細に示されているように、サーバ $S_X$ は、大容量または超大容量の唯一の仮想ディスク $D'_{XYZ}$ として、ディスクの集合 $D_X$ ないし $D_Z$ を“見る”。この場合、その容量は、少なくとも、ディスク $D_X$ ないし $D_Z$ の累積容量に等しくなる。必然的に、仮想ディスクの総容量は、サーバ $S_X$ の演算及び論理装置によってアドレス指定可能な最大ディスク空間に適合するように決められなければならない。結果的に、

サーバ $S_X$ は、あたかも上記の仮想ディスク $D'_{XYZ}$ ローカルディスクであるかのように、仮想ディスク $D'_{XYZ}$ のトラックまたはセクタ $S_m$ を直接的にアドレス指定する。このサーバは、上記の“I/Oチャネル”のジョブモードに従って、ローカルディスクに対するのと同じように、書込み及び読取りの基本的命令及び仮想ヘッドの移動命令を送る。

以下に詳述するように、本発明の好ましい実施形態においては、複数のタイプの“インテリジェントエージェント”が利用される。上記のアドレス指定が必要になることから、これらの特定エージェントを“記憶エージェント”と呼ぶことができる。物理装置 $DSM_X$ のRAMの中に、あるいは $DSM_X$ がサーバ $S_X$ の中に組込まれた論理ユニットである場合には、物理装置に代わってサーバ $S_X$ のRAMの中にロードされたプログラムの一部となる。

これらの基本的プログラムの役割は以下のように二つある。

a／例えば $D_X$ のようなローカルタイプであれ、例えば $D_Y$ 及び／又は $D_Z$ のように遠隔タイプであれ、アドレス指定された仮想トラックによる、特定の記憶装置への経路を設定すること。

b／実際の物理装置 $D_Y$ 及び $D_Z$ の機能に応じて、アドレス

指定モード及びプロトコルの翻訳を適合させること。

実際、本発明の右利な側面によれば、サーバ $S_X$ は、異質の記憶装置（異なるテクノロジー及び／又はメーカー、異なるジョブモード、等々）をアドレス指定することができる。例として、ディスク $D_X$ は、長さ32ビットの“ASCII”（American Standard Code for Information Interchange—情報交換用米国標準コード）に符号化されたワードを記憶するディスクとすることができ、サーバ $S_X$ は、“UNIX”オペレーションシステムの下で作動し、ディスク $D_Y$ 及び $D_Z$ は、長さ36ビットの“EBCDIC”（“Extended Binary Coded Decimal Interchange Code”）に符号化されたワードを記憶する“メインフレーム”と呼ばれる大出力コンピュータに結びついたディスクである。

物理装置またはその論理的等価ユニットは、サーバ $S_X$ の中に組込まれる場合には、物理的に情報を記憶するのがローカルディスク $D_X$ であれ、外部ディスク $D_Y$ または $D_Z$ であれ、ディスク $D'_{XYZ}$ のアドレス指定オペレーションを完全に透過性にする。より詳細には、この場合には、このタスクを遂行するの

は、記録された専用のインテリジェントエージェントあるいは記憶用インテリジェントエージェントである。サーバ $S_X$ は、あたかもそれが固有のローカルディスク $D_X$ であるかのように、サーバが“知っている”唯一またはいくつかのプロトコル、例えば、ローカルディスク $D_X$ に結びついたプロトコルに従って、仮想入／出力チャネル $I/O'_X$ を介して仮想ディスク $D'_{XYZ}$ をアドレス指定する。

第4図は、本発明に固有の配置（図では灰色の部分で示されている）を含むネットワークアーキテクチャ型情報処理システム1の一例を概略的に示している。

LANローカルネットワークは、“イントラネット”タイプであると仮定される。このタイプのネットワークは、ローカルに、“インターネット”ネットワークに固有のテクノロジーを組込んでいる。このネットワークは、略語“TCP/IP”で知られる基本的通信プロトコルを利用するサーバを備えている。このネットワークは更に、“インターネット”ネットワークでも同じように使用される

“HTTP”、“NFS”等々といった他の通信プロトコルを含んでいる。これらのプロトコルはすべて規格化されている。

更に、LANローカルネットワークは、例えば、以下に列挙するサーバのいずれか一つを介して“インターネット”ネットワーク  $I_T$  と通信すると仮定される。

説明した例においては、このLANローカルネットワークには、一方ではワークステーション  $ST_1$  ないし  $ST_3$  が、もう一方では記憶用サーバ  $S_1$  ないし  $S_3$  が接続される。これらの記憶用サーバ  $S_1$  ないし  $S_3$  を、簡略化するために、以下“サーバ”と呼ぶことにする。

本発明の主要な特徴一つによれば、各サーバ  $S_1$  ないし  $S_3$  に付与された記憶資源は、物理的または論理的な分散形管理装置  $DSM_1$  ないし  $DSM_3$  を介して、サーバに結びついている。ここに説明した例においては、従来のディスク用の二つの装置  $D_1$  ないし  $D_3$  が用いられる。

情報処理システム 1 は、磁気テープまたは磁気カートリッジのライブラリ  $TL_4$ 、一般に“ディスクファーム”と呼ばれるディスクの集合  $FD_6$ 、あるいはまた参照記号  $STO_e$  で示されている遠隔記憶資源のような他の記憶装置を備えることができる。これらの記憶資源はまた、それぞれに割当てられた管理装置  $DSM_4$ 、 $DSM_6$ 、 $DSM_5$  によって管理される。遠隔記

憶資源  $STO_e$  は、有利には、“ATM” (Asynchronous Transfer Mode - 非同期転送モード) モードにおいて回線  $I_{ATM}$  を介して、装置  $DSM_5$  と通信する。

本発明の他の側面によれば、さまざまな分散形管理装置  $DSM_1$  ないし  $DSM_6$  は、超高速バス  $B$  を介して互いに接続される。それらの装置はまた、同じくこのバス  $B$  を介して、集中形管理装置  $NSM$  に接続される。

例えば、“UNIX”オペレーションシステムの下で動作するワークステーションで構成される装置  $NSM$  は、データベース  $CDB$  を格納する、例えば一枚のディスクのようなメモリ手段を備えている。これらのデータは、システム、特に

あらゆる記憶資源及びそれらの特性（容量、ジョブモード、プロトコル等々）の記述を含んでいる。装置NSMはまた、情報処理システム1の各サーバS<sub>1</sub>ないしS<sub>3</sub>へのこれらの記憶資源の割当てを記述するデータを備えている。各記憶資源、より詳細には、磁気テープのライブラリTL<sub>4</sub>、“ディスクファーム”FD<sub>6</sub>、及び／あるいは遠隔記憶資源STO<sub>e</sub>は、複数のサーバS<sub>1</sub>ないしS<sub>3</sub>の間で共有することができることは明らかである。こうした共有は、一回だけ設定されるものではないという意味及

び／又は処理中のアプリケーションによって左右されるものであるという意味において、動的に行われる。磁気テープまたはカートリッジのライブラリは特に、システム1のディスク、一つまたは複数のサーバS<sub>1</sub>ないしS<sub>3</sub>に直接的にあるいは間接的に結びついたディスク上に記録されたデータの全部または一部の定期的な保護オペレーションに使用することができる。

記録されたデータベースCDBの上記のデータに基づいて、集中形管理装置NSMは、知られている方法で、上記の専用プログラムまたは“インテリジェントエージェント”を作成する。このため、本発明の有利な側面によれば、一方では“インターネット”ネットワークと結びついて使用される“J A V A - A p p l e t s”タイプのプログラミングテクノロジーが利用される。オブジェクト指向言語及び“実行時”タイプの環境、すなわち、そのプログラムがターゲット装置における受信の際に自動実行されることができる環境が使用される。実際に、これらプログラムの実行は、受信環境（“UNIX”オペレーションシステム、“Windows NT”“Windows 95”[商標]等々）によって左右されない。それは“Machine Virtuelle J A V A - J A V A 仮想装置”

と言われる。従って、“J A V A”アプリケーションは、ソフトウェア“Machine Virtuelle J A V A”がインストールされたあらゆる装置で実行することができる。すなわち、プログラムの実行は、使用されるプラットフォームとは無関係である。更に、通信は、“クライアントーサーバ”モードで

行われる。

本発明は、これらの有利な特徴を利用している。“インテリジェントエージェント”は、“J A V A”言語でプログラミングされ、高スループットバスBを介して、物理的であれ、論理的であれ（すなわち論理的ユニットの場合には、サーバと同一視される）、各分散形管理装置D S M<sub>1</sub>ないしD S M<sub>6</sub>に、動的にまた選択的に伝送される。受信においては、プログラムが装置D S M<sub>1</sub>ないしD S M<sub>6</sub>の中で自動的に実行される。この実行は、実際には、これらの装置（または、装置D S M<sub>1</sub>ないしD S M<sub>6</sub>が論理ユニットである場合にはサーバ）のR A Mへの命令のダウンロード及び格納である。

これらの命令は、特に、所与の瞬間に、所与の分散形管理装置（例えばD S M<sub>X</sub>）に結びついたサーバ（例えばS<sub>X</sub>）に、仮想メモリ空間（例えば第3 b 図：D' X Y Z）を割当ててゐる。

サーバ、例えばS<sub>1</sub>が、割当てられた仮想メモリ空間の中でデータの読取り及び／又は書込みのための要求を送信すると、分散形管理装置が、ダウンロード命令のコマンドの下に、ローカルディスクD<sub>1</sub>、もしくは例えばD<sub>3</sub>のような外部ディスクをアドレス指定する。いずれの場合にも、アドレス指定は、メッセージを用いて、通信プロトコルの形態においてではなく、“I / Oチャネル”モードで実行される。ローカルディスクD<sub>1</sub>については、アドレス指定プロトコルは、サーバS<sub>1</sub>によって使用されるプロトコルである。外部ディスクD<sub>1</sub>については、アドレス指定は、ディスクD<sub>3</sub>に結びついた分散形管理装置D S M<sub>3</sub>及びバスBを介して、この同じプロトコルによって実行される。この分散形管理装置D S M<sub>3</sub>は、使用されているハードウェアの異質性を補償するために、説明した方法でアドレス指定プロトコルを翻訳する必要がある場合もある。最終的なアドレス指定オペレーション、すなわち、物理ディスクD<sub>3</sub>の読取り－書込みヘッドのコマンドは、分散形管理装置D S M<sub>3</sub>の中にダウンロードされ記憶される命令のコマンドの下で実行される。これらの命令は、先の例と同じく、集中形管理装置N S Mによって伝送される“記憶用インテリジェントエージェント”

ト”のDSM<sub>3</sub>における自動実行から生成されるものである。

同じように、仮想ディスク空間における、サーバによる書込み及び／又は読取り要求は、チェーンの端で池の媒体、例えば磁気テープまたはカートリッジ上のデータの読取りまたは記録に物理的に翻訳される。これは、装置DSM<sub>4</sub>及び磁気テープライブラリTL<sub>4</sub>に到達する、例えばサーバS<sub>2</sub>によって出された要求の場合である。その結果生じるアドレス指定タイプの本質的変更は、サーバS<sub>3</sub>について透過性をもち続ける。必要な適応及びプロトコルの翻訳を実行するのは、DSM<sub>4</sub>にダウンロードされ記憶された命令である。この場合、DSM<sub>4</sub>はサーバに接続されていないので、論理的ではなく物理的な装置のみを対象とすることができることは明らかである。

システムが、本発明方法の必要条件を満たすことができるようになるためには、バスBが、上記の“I/Oチャンネル”タイプのプロトコルを受け入れる必要がある。換言すれば、バスにおける交換は、通信タイプのプロトコルに従っては、実行されない。

このために、本発明の一つの側面によれば、選択されるバスはANSI規格X3.230(1994)の対象である、い

わゆる“ファイバチャンネル”規格を満たすバスである。このタイプバスはまた、“ファイババックボーン”とも呼ばれる。それは、高速でデータを搬送することができるバスである。このようなバスに課される主な目的は、非常に短い待ち時間で、ある場所から他の場所へデータを伝送することにある。単純なエラーの訂正だけが、ソフトウェアでなく、ハードウェアによって行われる。データは二方向に同時に伝送される。混雑が原因で伝送が失敗しても、ソフトウェアの介入を必要とせずに、伝送をただちに再開することができる。更に、このタイプバスは、先に述べた“SCSI”のようなハイレベルのプロトコルと互換性がある。従って、このタイプバスは、“I/Oチャンネル”タイプの要求を搬送することができる。

バスには、三種類の形態が存在する。“ポイントツーポイント”タイプ、“Switched Fabric”と呼ばれるタイプ、そして“トークン型”リン

グに類似した“アービトレーションリング”タイプの三種類である。

本発明の範囲においては、できれば、最大の接続容量を有する第二の形態を選択することが好ましい。第5図は、このような形態を概略的に示している。この形態によれば、各装置、す

なわち本発明による分散形管理装置  $DSM_1$ 、 $\dots$ 、 $DSM_n$  ないし  $DSM_X$  及び集中形管理装置  $NSM$  の各々は、一つの交換器に接続されており、該交換器を介する他の接続のためのインターロックを有さないデータチャンネルを有する。この配置は、他の何からの装置との専用の接続に相当する。装置の数が増え、多数の交換器を占有すると、今度は、これらの交換器が互いに接続される。回路の冗長性を作り、総合通過周波数帯を大きくするためには、交換器間に多重接続チャンネルを確立することが好ましい。

規格は、リンクを形成するためにさまざまな物理媒体（撚り合わせケーブル、ビデオまたはミニチュア型同軸ケーブル、マルチモードまたはモノモードの光ファイバ）を使用することを許容しているが、本発明の範囲内では、モノモードファイバが選択される。このタイプの媒体は、超高速（ $100\text{ MB/s}$ ）であると同時に、長距離リンク（ $10\text{ km}$ まで）を可能にする。

従って、この選択は、 $100\text{ MB/s}$  までの速度について、典型的には  $10\text{ m}$  から  $10\text{ km}$  までのリンクを確立することを可能とし、この速度で数百の並列接続を受け入れることを可能にする。更に、バスBは、“TCP/IP”プロトコルのよう

な通信タイプであれ、または“I/Oチャンネル”（例えば“SCSI”）タイプのハイレベル型であれ、数多くのプロトコルの互換性を有する。バスBは、“I/Oチャンネル”タイプの多数のポートに相当する。

これにより、通常  $1/10$  程度に、二つのサーバ間のデータ交換に必要な時間を著しく短縮することができる。第6図は、“メインフレーム”タイプのコンピュータMF及び“UNIX”オペレーションシステムの下に作動するワークステーションSTXの二つのサーバ間のデータ交換を概略的に示している。本発明の

主な特徴によれば、ローカルハードディスク  $D_{MF}$  及び  $D_{STX}$  は、それぞれ、分散形管理装置  $DSM_{MF}$  及び  $DSM_{STX}$  を介してサーバに接続している。二つのディスクは、二つのサーバ間で完全に共有されると仮定する。従って、各サーバは、総ディスク空間を二つのディスクの総和に等しい容量の一つの仮想ディスク  $D_V$  として、認識する。

サーバの一方が要求を送信する、例えばサーバ  $MF$  が、仮想ディスク  $D_V$  におけるデータの書込みまたは読取り要求を送信するものと仮定する。この要求は、実行のために装置  $DSM_{MF}$  に伝送される。関連する物理ディスク空間が、ローカルディ

スク空間  $D_{MF}$  の中にない場合には、要求は、バス  $B$  を通して装置  $DSM_{STX}$  に伝送されることになる。この装置  $DSM_{STX}$  は、厳密な例においては、プロトコルの翻訳（先の例に示したように、データの記録モードは、それぞれ“ $EBCDIC$ ”と“ $ASCII$ ”の二つのディスクにおいて異なり、ワード幅も異なることから）と、伝送された仮想アドレスからのディスク  $D_{STX}$  の物理アドレス指定という二つのオペレーションを行なうことになる。

第7図のタイムチャートは、上記の交換に必要な時間を詳細に示している。総合時間間隔  $T'$  は、もはや三つのフェーズしかもたない。それぞれ、 $t=0$  及び  $t=t'_1$  の瞬間の間及び  $t=t'_2$  及び  $t=t'_3$  の瞬間の間の二つの両端の処理フェーズ  $T'_1$  及び  $T'_3$  と、更に、バス  $B$  を介した伝送フェーズである、 $t=t'_1$  及び  $t=t'_2$  の瞬間の間の  $T'_2$  である。

時間間隔  $T'_1$  及び  $T'_3$  と、第2図の時間間隔  $T'_1$  及び  $T'_5$ （従来の技術）とを比較しなければならない。先験的には、同じデータ処理／アプリケーション及びハードウェアについては、これら時間間隔は等しい。それらはまた、情報処理のローカルの条件にのみ依存し、伝送プロトコルにも伝送時間にも左

右されるものではない。

反対に、従来の技術とは逆に、時間間隔  $T'_2$  及び  $T'_1$ （第2図）といった通信プロトコルと関連した時間のロスはいもはや存在しない。これは、ハイレベルプ

ロトコル（例えば“SCSI”）との互換性をもつバスBの特性に起因する。更に、超高速バスと“Switched Fabric”と呼ばれる形態が使用されるので、伝送時間は最小限度まで短縮される。従って、時間間隔 $T'_2$ は、対応する時間間隔 $T_3$ （第2図）よりはるかに小さくなる。

こうして、一回の交換に必要な総時間間隔 $T'$ は、ほぼローカルの処理時間間隔にまで短縮され、実質的に、装置間の伝送から独立したものとなる。換言すれば、ローカル記憶資源とサーバとの間の交換、または外部記憶資源とこの同じサーバとの交換は、同一の速度でないとしても、少なくとも非常に接近した速度で行われる。従って、データの記憶資源の配置場所にかかわらず、システムの性能は低下しない。更に換言すれば、情報処理システム1のサーバ $S_1$ ないし $S_3$ のいずれも（第4図）、先験的に、性能を低下させる恐れなく、記憶資源 $D_1$ ないし $D_3$ 、 $TL_4$ 、 $FD_6$ または $STO_e$ の全てにアクセスすることが

できる。

結果的に、伝送時間（第6図： $T'_2$ ）が、許容可能な限度内に留まっている限りにおいて、情報のデータは、同一のサイトにおいてまたは遠隔サイトにおいて（第4図： $STO_e$ ）、情報処理システムのさまざまな記憶資源の中で、最適化に分配することができる。

上記の最適化された分配は、数多くの基準を考慮して行うことができる。例えば、さまざまな記憶資源の性能（速度）及び／又は容量、データのタイプ、全体または部分的な、一つまたは複数の記憶資源の一時的故障、容量オーバー、記憶コスト等々といった基準である。

すでに示したように、分散形記憶管理は、データベースCDBを考慮して、集中形管理装置によって行われる。このデータベースCDBは、さまざまな分散形管理装置 $DSM_1$ ないし $DSM_6$ の中で、記憶インテリジェントエージェントをダイナミックに作成し、ダウンロードする。このタイプのエージェントは、特に、ローカルの条件（接続されたハードウェアのタイプ、ローカルメモリの容量、使用されるプロトコル等々）と、任意のサーバに割当てられた総仮想ディスク空間と、一つまたは複

数の物理記憶資源を管理する分散形管理装置または集中形管理装置のアドレスを計算に入れる命令を搬送する。これらの特性は、時間に応じて変化する。更に、エージェントのコマンドの下で行われたオペレーションはダイナミックである。

実質的に、従来のデータ伝送ローカルネットワークの管理と類似の方法で、管理が行われる。集中形管理装置NSMは、分散形管理装置DSM<sub>1</sub>ないしDSM<sub>6</sub>の“ネットワーク”と呼ぶことができることから、“UNIX”におけるサーバのようなデータの処理装置によって具体的に表わされ、更に、バスBと共存する従来のLANローカルネットワークの管理装置と同一視することができる。

記憶インテリジェントエージェントに加えて、本発明の範囲内で、他のタイプのエージェントを有効に利用することができる。

第一のタイプの追加エージェントは、専用の入／出力オペレーションの管理エージェントによって構成される。これらのエージェントは、上記の入出力オペレーションについて特有の結果を得るために、特定の記憶ノードの中に、すなわち、実質的にこれらノードに結びついた分散形管理装置の中にダウンロード

ドされる。これは例えば、データベースへのアクセス時間の最適化を可能にする。

第二のタイプの追加エージェントは、一定数のパラメータ、例えば日付、時刻、データのタイプ等々に応じて、自動的バックアップを組織するエージェントによって構成される。これらのエージェントは、所定のデータの記憶手段に向けて、複製によるデータの移動を組織する。

第三のタイプの追加エージェントは、システムの中に存在する記憶資源間のデータの階層的記憶及び保存を管理するエージェントによって構成される。これらのエージェントは、特に、アクセス時間、アクセス頻度、ローカリゼーション及びコストのような制約条件を考慮して、一つの資源から他の資源へデータを移動することができる。より詳細には、特有のエージェントは、“メディアのサーバエージェント”と呼ぶことができる磁気テープまたはカートリッジのライブラリTL<sub>4</sub>に結びついた分散形管理装置に割当てられる。このエージェントは、物理記憶オブジェクトの仮想ビジョンをもたらし、外部サーバからの入／出力要求に

合わせることができる。これら外部サーバは、あたかも、例えばディスクのようなローカルの記憶資源を処理

しているかのように、従来の読取り-書込み要求を送信しながら、これらの記憶オブジェクトにアクセスする。この手段によって、アプリケーションのサーバは、アプリケーションを変更することなく、現在実用に供されているものであれ、開発中のものであれ、大容量記憶用の何らかのテクノロジーを有効に利用することができる。

第四のタイプのエージェントは、例えば、データのプロファイル、特有のアプリケーション及び発生するコストに応じて、冗長記録手段を管理するエージェントによって構成される。安全性、特にデータの無傷性と使用可能性の理由により、データは、多少とも増強された冗長性ととも記録され、更に複製される。また、情報の有効なデータに対して、パリティキー、エラー検出コード（“Error-Detection Code”の略で“EDC”）またはエラー補正コード（“Error-Correcting Code”の略で“ECC”）といった冗長データが付け加えられる。このような構成から、“ミラー”または“RAID 5”と呼ばれるディスク、すなわちパリティキーの計算をとまなうディスクのようなさまざまな記録技術が導かれる。この第四のタイプのエージェントは、ある技

術または他の技術の使用が最適化されるように管理する。

第五タイプのエージェントは、データの記憶用の遠隔サイトSTO<sub>e</sub>との接続を管理するエージェントで構成される。すでに示したように、リンクは“ATM”モードで行われる。従って、“ファイバチャネル”から“ATM”へのプロトコルの翻訳を行なうことが必要である。このタイプのエージェントは、他のエージェント（記憶エージェント等々）に加えて、分散形管理装置DSM<sub>5</sub>の中にダウンロードされる。

第六のタイプのエージェントは、“インターネット”ネットワークとの交換を管理し、この手段によって、情報処理システム1の一つまたは複数の記憶資源に

向かって、ダウンロードされたデータを経路設定するために、分散形管理装置のいずれかの中でダウンロードすることができる。

第七のタイプのエージェントは、データの記憶資源の集合の正確なジョブを監視及び管理するエージェントによって構成される。

他のタイプのエージェントは、必要に応じて、集中形管理装置NSMによって構成され、バスBを介して分散形管理装置の中に選択的に遠隔ダウンロードすることができる。このような

エージェントは特に、第一のプロトコルを第二のプロトコルに翻訳することが必要となる場合に必要となる。

同様に、データ処理の安全性に関し、秘密性が考慮されなければならない。この点はまた、専用のエージェントを利用しながら扱うことができる。これらのエージェントは、例えば、要求が、データの内容、ユーザープロファイルまたはアクセスされた資源に応じて、正当である及び／又は予約されているかどうかをチェックすることができる。

第4図に示されているアーキテクチャには、二つのバスが共存している。リング型のローカルバス（記載例においては）LANと、本発明に特有のバスである高速バスBである。このアーキテクチャによって、記憶サーバ $S_1$ ないし $S_3$ と、他の装置、例えばワークステーション $ST_1$ ないし $ST_3$ との間の通信は、LANローカルネットワークのタイプと互換性をもつ通信プロトコルを利用しながら、ファイルの転送の形態で従来の方法で行われる。

しかしながら、“ファイバチャネル”スタンダードに適合したバスBは、通信タイプのプロトコルを含む複数のプロトコルを搬送することができるので、LANネットワーク（第4図）

を排し、あらゆる伝送が唯一のバスBを通過するようにさせることが可能となる。

第8図は、唯一つのタイプのバス、バスBのみを備えている情報処理システム1'のアーキテクチャを概略的に示している。あらゆる装置は、本発明に固有の

ものであろうとなかろうと、このバスBに接続される。例としての、この第8図には、分散形管理装置DSM<sub>1</sub>を介してローカルディスクD<sub>1</sub>に結び付けられたサーバS<sub>1</sub>と、分散形管理装置DSM<sub>7</sub>を介してバスBに接続され、“インターネット”ネットワークI<sub>T</sub>との通信を管理する第二のサーバS<sub>7</sub>と、磁気テープまたはカートリッジのライブラリTL<sub>4</sub>に結びつけられた分散形管理装置DSM<sub>4</sub>と、集中形管理装置NSMと、ディスクCDBとして表わされたそのデータベースと、更に、後にバスBに直接接続される三つのワークステーションST<sub>1</sub>ないしST<sub>3</sub>が示されている。ワークステーションST<sub>1</sub>ないしST<sub>3</sub>は、従来の方法で、すなわち、通信プロトコルを使用して、互いに通信する。実際に、すでに示したように、二つのタイプのプロトコルは、バスB上に共存することができる。この意味において、サーバ、例えば、サーバS<sub>1</sub>は、バスBに直接接続され、例えば“TC/IP”

といった通信プロトコルの下で他の装置（例えばST<sub>1</sub>ないしST<sub>3</sub>）と通信することができる。

上記から、本発明はその目的を達成することができることが容易に理解される。

特に、本発明は、各サーバに対して、総合的記憶空間の仮想空間を提供することができる。先験的に、本発明によって、いかなるサーバも、その性質に関わらずシステムのどの記憶資源にもアクセスすることが可能になる。更に、本発明は、性能を大きく低下させることなく、遠隔の資源にアクセスすることを可能にする。

しかしながら、本発明は特に、第3図から第8図に関連して詳しく説明した実施例のみに限定されるものではない。本発明は数多くのタイプのハードウェアに適合する。まず第一に、これは、マイクロコンピュータ、ワークステーション、ミニコンピュータ、または“メインフレーム”と呼ばれるタイプの大型コンピュータを備えることができるデータ処理中央装置についてあてはまる。オペレーションシステムもまた多種多様とすることができ、例えば汎用型または所有権と呼ばれるオペレーションシステムを備えることができる。バスBに加えて、バス（

例

例えば“イーサネット”タイプ)あるいはリング型(例えば“トークンリング”タイプ)といった多種多様なローカルネットワークを使用することもできる。更に、また特に、記憶用周辺装置のタイプも限定されない。すなわち、ローカルサイトまたは遠隔サイトにおいて、磁気ディスク、光ディスク、磁気テープ及び／又はカートリッジを利用することができる。

【図1】

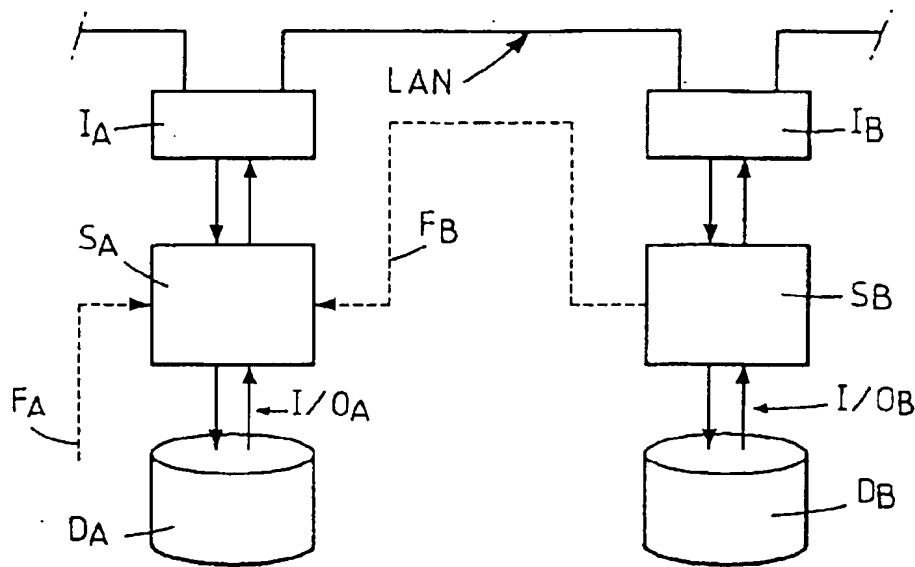


FIG.1

【図2】

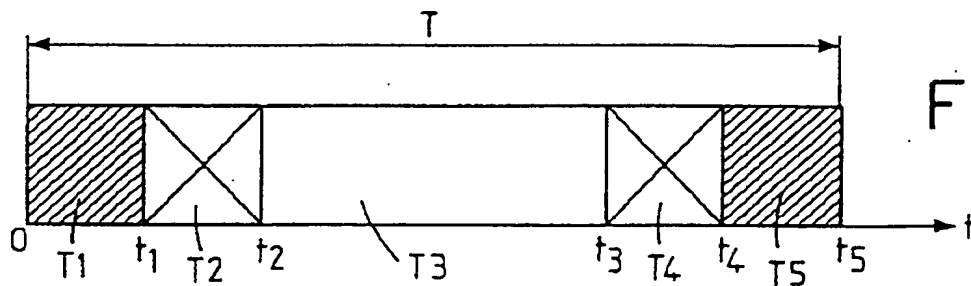
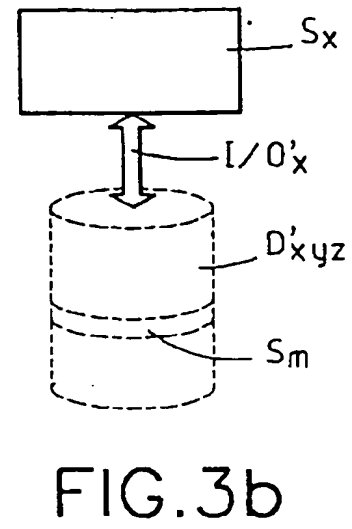
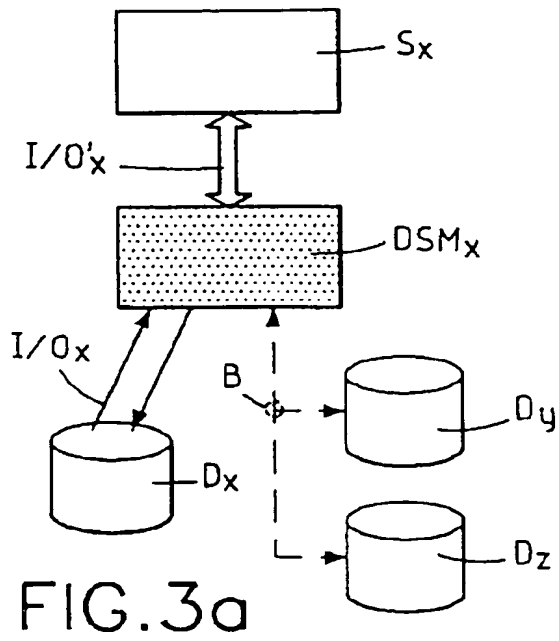


FIG.2

【図 3】



【図4】

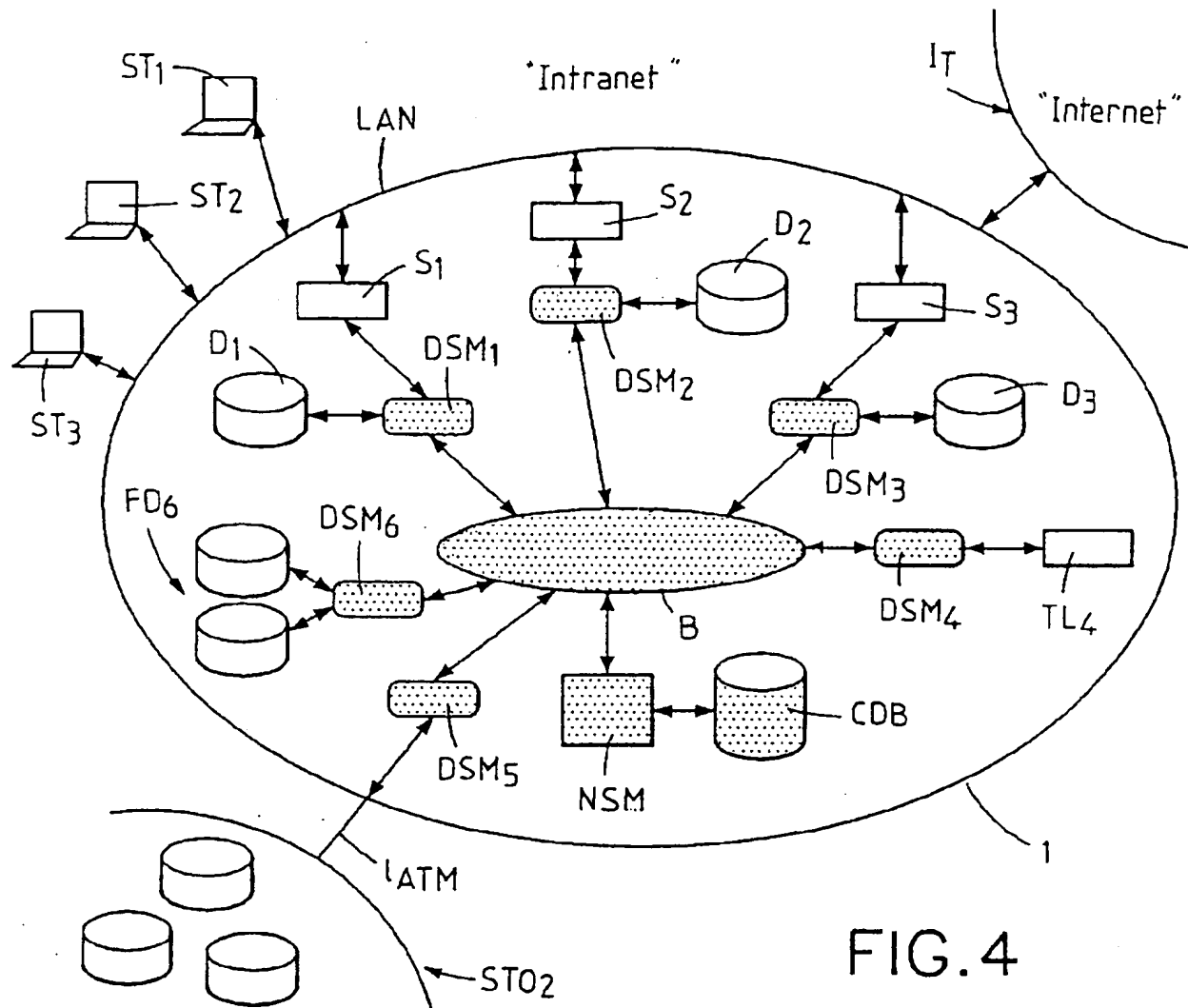


FIG. 4

【図5】

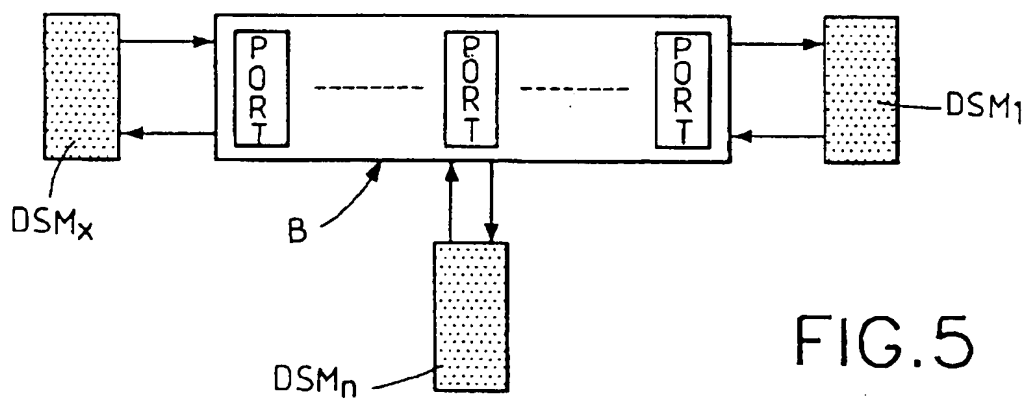


FIG. 5

【図 6】

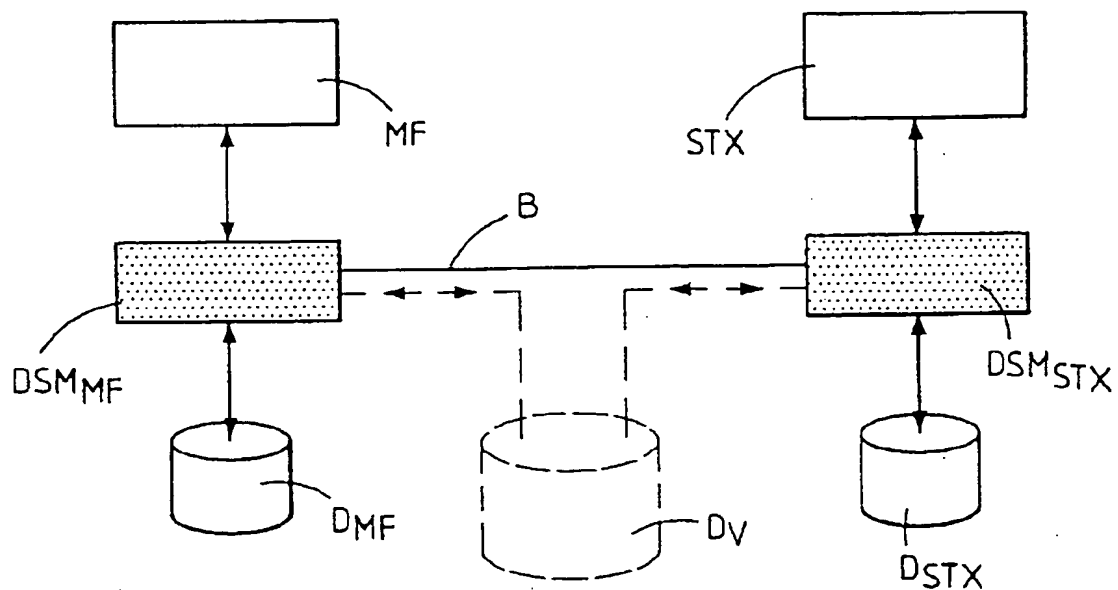


FIG. 6

【図 7】

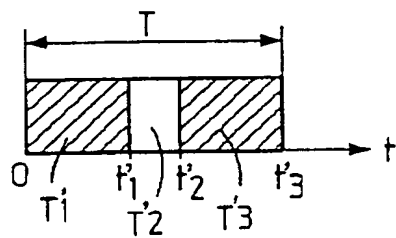
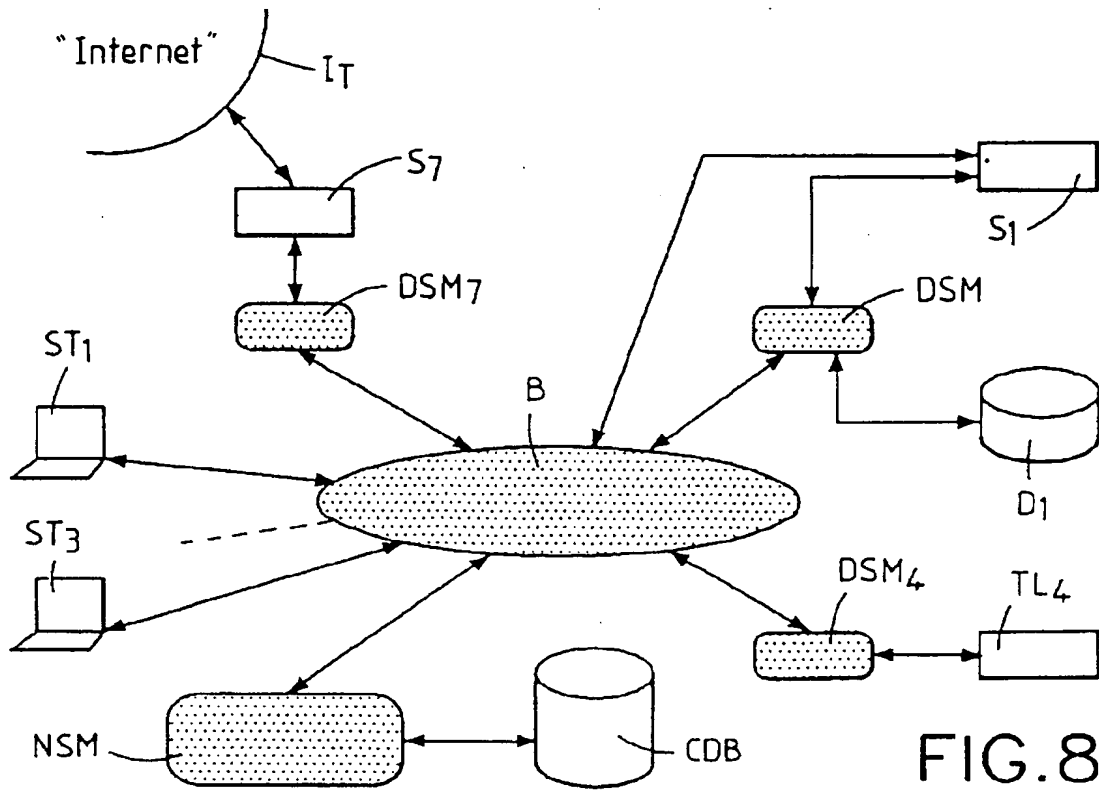


FIG. 7

【図8】



【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International Application No.

PCT/FR 98/00110

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	WO 89 02631 A (DIGITAL EQUIPMENT CORP) 23 March 1989	1,2
Y	see abstract see page 4, line 19 - page 5, line 7	7,16
A	EP 0 747 840 A (IBM) 11 December 1996 see abstract	1-17
Y	see column 2, line 10 - line 31 see column 5, line 13 - line 41 see column 6, line 59 - column 7, line 7	16
X	US 5 367 698 A (WEBBER NEIL F ET AL) 22 November 1994	1,2,15
Y	see abstract see column 1 - column 3	7

-/--



Further documents are listed in the continuation of box C



Patent family members are listed in annex

## \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited in understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"S" document member of the same patent family

Date of the actual completion of the international search

14 July 1998

Date of mailing of the international search report

24/07/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epc nl,  
Fax: (+31-70) 340-3016

Authorized officer

Adkhis, F

## INTERNATIONAL SEARCH REPORT

Internat. Application No.

PCT/FR 98/00110

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	EP 0 774 723 A (MATSUSHITA ELECTRIC IND CO LTD) 21 May 1997 see abstract see column 3, line 37 - column 4, line 12 -----	1, 2

## INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern: .al Application No

PCT/FR 98/00110

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 8902631 A	23-03-1989	CA 1312385 A DE 3889904 D DE 3889904 T EP 0338041 A JP 1502786 T US 5408619 A	05-01-1993 07-07-1994 12-01-1995 25-10-1989 21-09-1989 18-04-1995
EP 0747840 A	11-12-1996	US 5752246 A CN 1143776 A JP 9026973 A	12-05-1998 26-02-1997 28-01-1997
US 5367698 A	22-11-1994	NONE	
EP 0774723 A	21-05-1997	JP 10003421 A	06-01-1998